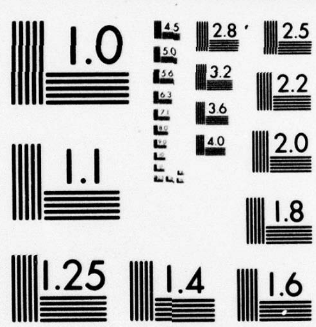
AD-A067 765 AIR FORCE HUMAN RESOURCES LAB BROOKS AFB TEX F/G 5/9
F-15 FLIGHT SIMULATOR: DEVELOPMENT AND ANALYSIS OF COMPUTER SCO--ETC(U)
MAR 79 M J MCDONALD, B A SMITH, D W EVANS

UNCLASSIFIED AFHRL-TR-78-72 NL

1 OF 1
AD
A067765

END
DATE
FILMED
6-79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFHRL-TR-78-72

LEVEL II

AIR FORCE HUMAN RESOURCES LABORATORY

F-15 FLIGHT SIMULATOR:

DEVELOPMENT AND ANALYSIS OF COMPUTER SCORING ALGORITHM

By

Michael J. McDonald, 2nd Lt, USAF
Bruce A. Smith, Capt, USAF
David W. Evans, 2nd Lt, USAF

USAF ACADEMY, Colorado 80841

Lester H. Baer, Maj, USAF

TACTICAL AIR WARFARE CENTER
Luke Air Force Base, Arizona 85309

William H. Nelson

FLYING TRAINING DIVISION
Williams Air Force Base, Arizona 85224

March 1979
Final Report for Period June 1977 – September 1977

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

When U.S. Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This final report was submitted by Flying Training Division, Air Force Human Resources Laboratory, Williams Air Force Base, Arizona 85224, under project 1123, with HQ Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235. Dr. William Nelson (FTO) was the Principal Investigator for the Laboratory.

This report has been reviewed and cleared for open publication and/or public release by the appropriate Office of Information (OI) in accordance with AFR 190-17 and DoDD 5230.9. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved for publication.

DIRK C. PRATHER, Lt Col, USAF
Technical Advisor, Flying Training Division

RONALD W. TERRY, Colonel, USAF
Commander

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFHRL-TR-78-72	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) F-15 FLIGHT SIMULATOR: DEVELOPMENT AND ANALYSIS OF COMPUTER SCORING ALGORITHM		5. TYPE OF REPORT & PERIOD COVERED Final rept. June 1977 – September 1977
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Michael J. McDonald, Lester H. Baer, Bruce A. Smith, William H. Nelson, David W. Evans		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Flying Training Division Air Force Human Resources Laboratory Williams Air Force Base, Arizona 85224		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62205F 11231209
11. CONTROLLING OFFICE NAME AND ADDRESS HQ Air Force Human Resources Laboratory (AFSC) Brooks Air Force Base, Texas 78235		12. REPORT DATE March 1979
		13. NUMBER OF PAGES 30
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) 30p.		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

computer scoring algorithm
flight simulator
flying training research
objective standards
operational test and evaluation
performance measurement
pilot training

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This study was designed to develop and evaluate the computer scoring algorithm of the F-15 flight simulator. Subjects were F-15 pilots in the grade of 1st Lt through Lt Col with previous flying experience commensurate with grade and operational assignments. Evaluation involved simultaneous scoring by the computer and Instructor Pilots (IP) of flight departures and approaches. Both scores were then compared to estimate the validity of the computer algorithm. Departure scores were moderately correlated (r = .75); however, approach scores exhibited moderate to high negative correlation (r = .01 to –.91). Interaction from the IPs indicated that scoring parameters measured by the computer were correct. It was determined that the negative correlations on the approaches were a result of the computer initiating scoring whenever a certain range boundary was reached, whereas the IPs began scoring only

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

404 415

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Item 20 Continued:

when the appropriate legs of the approaches were being flown. This variance is being investigated and a modification has been recommended. Because of the correlation of the departure scores, it was concluded that with improvements to the computer scoring procedures for the approaches, the scoring algorithms of the F-15 flight simulator could provide a valuable tool for evaluation of fighter pilots.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

EXECUTIVE SUMMARY

Objectives

The primary objectives of this study were (a) to evaluate student scoring algorithms, including parameters, in the F-15 Flight Simulator (FS), and (b) to recommend adjustment of parameters scored and parameter tolerances, accordingly.

Approach

In this study, comparison between the FS automatic computer scoring and the instructor pilot (IP) scoring were made on one departure and three approach flight patterns used at Luke AFB. Each pattern was made up from a number of leg segments in sequential order. The parameter values for the various leg segments of these flight patterns were defined by the pattern restrictions for them, e.g., terrain clearance, congested areas of population, and altitude restrictions. The tolerances for these flight patterns were defined as flight check criteria in Tactical Air Command (TAC) Regulation 60–2.

Although the FS computer has the capability of recognizing 15 parameters for evaluation and scoring, it can only grade three of the parameters on each flight leg segment. A choice must be made as to which three of the parameters are most useful for evaluation within their acceptable tolerances for each leg segment of the pattern.

Background

In April 1977, TAC began an Operational Test and Evaluation (OT&E) of the F-15 FS, built by Goodyear Aerospace Division and physically located at Luke AFB. The Air Force Human Resources Laboratory, Flying Training Division, Tactical Research Branch (AFHRL/FTO) at Luke AFB assisted in the conduct of several areas of the OT&E.

A major component of the FS system is the Instructor Operator Station. This station is so designed that the IP can instruct the student in the cockpit, observing his progress both through instruments and cathode-ray tubes on the console, and can initiate a variety of tactical problems for the student. It is from this location that the IP also grades the student on his performance.

Since the FS has the capability of objectively scoring flight performance and since it was designed to be an aid to the IP-student relationship, the computer-derived evaluations must agree with the IP evaluations of performance. This study was designed to satisfy one objective of the OT&E– i.e., evaluate the Student Scoring Algorithms, including Parameters.

Specifics

The subjects in this study were all Air Force pilots undergoing the F-15 training program at Luke AFB. Pilots were from Nellis AFB and Luke AFB (the 555th, 461st, and 433rd TFTS). Each unit provided the IP for its own students. Each subject received twelve periods of instruction, each lasting 1-1/2 hours, for a total of 18 hours of instruction. The IP was briefed prior to the start of each instructional period as to the purpose of the study and what was desired of him. His completed grade forms were reviewed after each period by an investigator.

The IP evaluations of the student's performance were made on a modified Air Force flight grading form. This modification permitted each leg segment to be graded on a continuum from 0 to 4 with an expanded middle range from 1 and 3. This expansion (80 percent of the total range) permitted more sensitive discrimination in this range. The IPs were instructed to place a mark anywhere on the continuum that corresponded to their evalautions.

Both the departure and approach flight patterns were divided into leg segments in such a manner that the computer and the IP graded the same component parts of the pattern. The computer scoring algorithms

were designed to grade amounts of deviation from the acceptable flightpath tolerances as well as the time involved during the deviation(s). Points were subtracted from 100 for each deviation, with –99 being the lowest possible score.

The original computer scoring algorithm had not been operationally tested, and following the simulator period of instruction, the IPs were asked to select what they felt to be the two most important parameters for each leg. The design of the experiment was such that the nature of the expanded grading scale approximated interval data, thus permitting analysis between the computer scoring algorithms and the IP evaluations with both a Pearson Product Moment Correlation and a Spearman Rank Coefficient statistic. There was very good agreement between the two methods of grading on the one departure flight pattern, with significant positive correlations (.64 to .75) obtained. However, on the approach flight patterns, little or negative agreement was found to exist between the two methods of grading.

It became apparent from the analysis and inspection of data that the major problem was intrinsic to the computer. The departure flight pattern was always started from one specific point in space–the runway–and both the computer and the IP initiated scoring at identical points in time and space. Not all students, however, entered the approach flight patterns at the same point in space and time. It was determined that the poor agreement between the two scoring methods on the approach flight patterns was due to the computer initiating its scoring whenever a certain range boundary was reached by the student, even though the student might depart from that point, exit the boundary, and return at another point on the boundary whereas the Instructor Pilots began their scoring only when the student began flying the appropriate leg segments of the approach flight pattern. This variance is being investigated and proposals have been recommended to alleviate this problem.

The differences found between the parameters selected by the IPs and the parameters currently in use by the computer scoring algorithms were found to be negligible for three of the four flight pattern profiles. The IPs suggested changes in 28 percent of the possible parameters for one approach flight pattern.

Conclusions

Through the use of scoring algorithms, the computer can give an unbiased evaluation of one student's performance against that of another student on identical tasks and task sequences. The computer can also give an unbiased evaluation of a student's performance against established standards. It cannot, however, take into account other factors considered by the IP in grading overall performance, such as airmanship, safety of flight, or actual flying techniques.

The FS is a valuable training asset and, with improvement, will greatly enhance flying training and the quality of such training.

PREFACE

This study was conducted in support of Tactical Air Command, (TAC) Project 75A-034U Operational Test and Evaluation (OT&E) of the F-15 Flight Simulator; Maj Lester H. Baer, Project Officer. This report covers research conducted between June and September 1977.

This research was conducted by Capt Bruce A. Smith and Cadets Michael J. McDonald and David W. Evans of the United States Air Force Academy sponsored by Frank J. Seiler Research Laboratory, as a summer research project under the direction of Dr. William Nelson, Air Force Human Resources Laboratory (AFHRL/FTO) at Luke AFB.

Special appreciation is expressed to the research support personnel at Luke AFB and Williams AFB for their assistance; TAWC/OLAH, commanded by Col Ray F. McNally, for their support; the airmen and non-commissioned officers of the 58th Component Repair Squadron for their assistance and professionalism; to AFHRL/FTO headed by Mr. Robert Bunker; and to Mr. Robert Coward, Senior Research Psychologist of AFHRL/FTO for his help throughout this project.

TABLE OF CONTENTS

LIST OF TABLES

F-15 SIMULATOR: DEVELOPMENT AND ANALYSIS OF COMPUTER SCORING ALGORITHM

I. INTRODUCTION

Background and Purpose

Pilot performance is obviously a subject of great interest in the Air Force. The whole concept of pilot performance assumes that there are differences in pilots and that there are measurable characteristics that distinguish these differences (Fessler, Note 1). Two major schools of thought are concerned with how to measure these differences. One, the traditional method, emphasizes the check pilot evaluator using completely subjective measures. The second emphasizes the factual, non-emotional, machine-type scoring, using objective measures which are directly applicable to simulator flying (Fessler, Note 1).

Research done in the subjective methods of scoring has shown that a number of difficulties exist. Although the human analytical approach has the advantages of human involvement and fully exploits human understanding and observational powers, it is time-consuming (Knoop, 1973). Koonce (1975) and Obermayer, Vreuls, Muckler, Conway, and Fitzgerald (1974) noted the importance of providing the human observer with objective standards. The importance of objective measures for pilot prediction was pointed out by Johnson and Boots (1943) during World War II and was further emphasized by Mahler and Bennett (1950) and Ericksen (1951). The importance of setting a baseline for comparative measurement was stressed by those investigators. Regardless of how carefully designed and controlled a subjective measurement is, it would still experience some difficulty in lack of reliability and in inconsistent discrimination (Fessler, Note 1).

With this in mind, the development of objective systems seems to be appropriate in order to obtain the most advantagenous method of scoring. This leads to design philosophies that stress reliance on behavioral task analysis data (Cream & Lambertson, 1975). There are, however, problems existing in these areas. One of the most critical problems is that the maneuvers which seem to be the best discriminators of performance are also extremely difficult to measure. Although the machines in the early 1970's were somewhat below the technological level needed for scoring these maneuvers (Long & Verney, 1976), this problem has been alleviated by the advent of the modern, digital-computer flight simulators, such as the Advanced Simulator for Pilot Training (ASPT), the Simulator for Air-to-Air Combat (SAAC), and the F-15 Flight Simulator (FS). Because of the special capabilities of these new generation simulators, the concept of objective measurement is becoming a reality. It is necessary, however, to insure that the measurement criterion, baseline measures, and resulting performance scores are validated before applying them to operational situations.

The criterion for evaluating any flight simulation device is its training effectiveness (Waag, Eddowes, Fuller, & Fuller, 1975). Flight simulators are designed to instruct students through safer, cheaper, and possibly more effective means than those available in actual flight conditions. Flying is criterion-directed, however, and maneuvers in the simulator must have "...definable objectives which must be accomplished" (Waag et al., 1975). These objectives acquire extreme importance and must be able to be measured in terms of parameters available to the simulator's scoring computer. Waag et al., 1975, suggests two constraints of such parameters: (a) the measures will assess the degree to which the criterion objectives are met and (b) the measures will reflect only the most salient characteristics of performance. Such an approach was used in this study to evaluate the use of defined parameters for three approaches–Luke HI-ILS RWY 3R, ILS RWY 11, and the JAY HI-TACAN; and one departure, the FLATIRON Departure 03 (see Appendix A).

The evaluation of performance for a given parameter involves a comparison of the obtained value with some ideal value" (Waag et al., 1975). In this study, the ideal value for the approaches and departures

was defined by the pattern restrictions for the maneuvers, and the tolerances were defined as flight check criteria in Tactical Air Command (TAC) Regulation 60–2. However, although the F-15 FS recognizes 15 parameters of evaluation, it is capable of grading only three per flight leg segment. A choice, therefore, must be made as to which parameters are to be evaluated and the acceptable tolerances.

Since the F-15 FS was designed to be an aid to the instructor pilot (IP)–student relationship, it is necessary for computer evaluations to correlate with IP evaluations of the student's performance. In this light, measures must not only assess the degree to which objectives are met and the most salient characteristics of performance, but do so in such a manner that they agree with IP interpretation of the maneuver. This study was to fulfill that requirement, and to recommend adjustments in parameter scoring and parameter tolerances in order to improve correlations between IP and computer scores.

II. METHOD

In April 1977, TAC began an operational test and evaluation (OT&E) of the F-15 FS, built by Goodyear Aerospace Division and located at Luke AFB. Air Force Human Resources Laboratory, Flying Training Division, Tactical Research Branch (AFHRL/FTO) at Luke AFB, was asked to assist in the conduct of the operational evaluation of the F-15 FS in several areas. This report deals with Objective #3 of the OT&E – Evaluate the Student Scoring Algorithms, Including Parameters.

Subjects

The subjects in this study were all Air Force pilots from Nellis AFB and Luke AFB (555th, 461st, and 433rd TFTS) in the F-15 training program at Luke AFB. Each unit provided the IP for its own subjects. The subjects ranged in grade from O–2 to O–5 and their experience in aircraft and simulator hours were commensurate with grade and operational assignments. The IPs were more uniform in grade, clustering around junior to senior captain. The subjects utilized were those undergoing training in the F-15 simulator at Luke AFB from 6 June to 8 August 1977, which involved a total of six classes of students. The IPs remained constant for each unit.

Apparatus

The F-15 Flight Simulator at Luke AFB was the experimental device for study and analysis. The simulator has a six-degrees-of-freedom motion base; however, it does not have a visual field other than the Heads-Up Display. It was designed by the Goodyear Aerospace Division and was maintained jointly by Goodyear and Air Force personnel.

The simulator had been programmed to evaluate the subject's performance in certain areas of flight. The performance of the subject in the simulator was graded both by the IP and by the computer, providing a basis for comparison. One departure from and three approaches to Luke AFB and its auxiliary field were selected for evaluation: the FLATIRON 03 departure and the LUKE HI-ILS 3R, ILS RW 11, and JAY HI-TACAN approaches. Although the computer recognizes 15 parameters, only three may be used on any one scoring leg segment. The one departure and three approaches have been broken down into legs (See Appendixes A and C). The scoring algorithm grades the deviation from and time out of the accepted standards. Points are then subtracted from 100, with –99 being the lowest possible grade. In brief, if the student exceeds the tolerance of any scored parameter, he is penalized points according to the following equation:

1. Out of tolerance on any monitored parameter – each time	N x 1
2. Fifty percent over tolerance – each time	N x 2
3. One hundred percent over tolerance – each time	N x 3
4. Selecting incorrect TACAN channel	N x 1

Where N equals the number or parts of 30-second intervals in each out of tolerance, and the score is 100 minus the sum of penalties above (see Appendix D).

One must recognize that the alert IP considers other factors in overall scoring, such as airmanship, safety, and actual flying techniques; however, the computer score gives an actual unbiased evaluation of one subject's performance against that of another on identical tasks and task sequences.

Instructor evaluations were completed on a modified AF Form 1363 that was reidentified as special AFHRL Form 135 (shown in Appendix A). The departure and approaches were divided into the same subparts utilized by the computer. The IP evaluation of the subject's performance on each leg was marked on a continuum from 0 to 4 with an expanded middle range.

The IPs were familiar with a 0 to 4 scale range, since they used it for standard mission grading on AF Form 1363. To avoid confusion and remain compatible with ongoing instruction, this study also utilized the 0 to 4 range. The criteria definitions of the grading remained the same; however, to avoid the error of central tendency, and yet obtain more sensitive discrimination, the range from 1 to 3 was expanded as shown in Appendix A to include 80 percent of the scoring range. The IPs were instructed to place a mark anywhere on the continuum that corresponded to their evaluation. Following the simulator period of instruction, the IPs indicated what they felt to be the two most important parameters for each leg from a list of eight parameters. These eight were based on their repeated appearance throughout the initial phases of this research (see Appendix A).

Because of the small number of responses and to allow more time for reflection, the IPs were later asked to again select parameters on a separate questionnaire. The legs of each graded portion of the flight were described and then presented visually on a map (see Appendix B for sample). The IPs checked the three most important parameters for each leg and were encouraged to add one parameter that they considered critical from a numbered list provided. The first questionnaire was given to and completed by the IP in the F-15 simulator building. The later form was completed in the squadron area. The rationale for this is discussed later in the text.

Procedure

Since the analysis was to be done on the effectiveness of the existing system, the initial parameters and standards of acceptance were presumed to be correct. A program of the algorithm was obtained and compared with flight maps to determine the description of each leg.

Each class of subjects had 12 periods of instruction, each lasting 1–1/2 hours. The IP was briefed at the beginning of each period on the purpose of the study and what would be required of him. The experimenter reviewed the completed grade form after each period.

The first questionnaire section, on solution of parameters, was often completed hastily due to a local requirement to complete the actions necessary during the scheduled simulator time. Therefore, a second questionnaire was prepared and distributed to the squadrons through their operations desk approximately 2 weeks later (see Appendix B).

An average of the performance evaluations was calculated for each approach or departure. These were ranked and matched with their respective computer score. As provided in the research design, a Pearson Product Moment Correlation and a Spearman Rank Coefficient analysis were then performed on the data. The expanded grading scale approximated the interval data, therefore, allowing for analysis with the Pearson's correlation as well as with the Spearman's. The results of the parameter questioning on both the first and second surveys are presented in Tables 1 to 4.

Table 1. **Comparison of Algorithm and Instructor Parameter Measurement (JAY HI-TACAN APPROACH)**

Mission Element Number	Parameters Currently Programmed	Parameters Selected by IPs	Additions/Corrections
1	Altitude, Channel	Altitude, Channel, Airspeed	Airspeed
2	Airspeed, Altitude, Radial[a]	Airspeed, Altitude, Radial[a]	
3	Airspeed, Radial[a]	Airspeed, Radial,[a] Configuration	Configuration
4	Airspeed, Altitude	Airspeed, Altitude, DME	DME
5	Airspeed, Altitude, DME	Airspeed, Altitude, DME	
6	Airspeed, Altitude	Airspeed, Altitude, Radial[a]	Radial[a]
7	Airspeed, Altitude, Radial[a]	Airspeed, Altitude, Radial[a]	
8	Airspeed, Altitude, Configuration	Airspeed, Altitude, Configuration	
9	Airspeed, Altitude, Radial[a]	Airspeed, Altitude, Radial[a]	
10	Airspeed, Altitude, Radial[a]	Airspeed, Altitude, Radial[a]	
11	Airspeed, Altitude, Radial[a]	Airspeed, Altitude, Radial[a]	
12	Airspeed, Altitude, Radial[a]	Airspeed, Altitude, Radial[a]	

[a]Radial includes Heading Information.

Table 2. **Comparison of Algorithm and Instructor Parameter Measurement (ILS RUNWAY 11 APPROACH)**

Mission Element Number	Parameters Currently Programmed	Parameters Selected by IPs	Additions/Corrections
1	Airspeed, Altitude, Channel	Airspeed, Altitude, Heading	Heading
2	Localizer Deviation (LOD)[a], Altitude	LOD[a], Altitude, Heading	
3	Airspeed, LOD[a], Configuration	Airspeed, Altitude, Radial	Altitude
4	Airspeed, Glide Slope Deviation (GSD)[b], LOD[a]	Airspeed, Altitude, Radial	
5	Altitude, Glide Slope Deviation[b], LOD[a]	Airspeed, Altitude, Radial	Airspeed
6	LOD[a]	Airspeed, Altitude, Radial	Airspeed, Altitude

[a]LOD includes both Radial and Heading Information.

[b]Glide Slope Deviation includes Altitude Information.

Table 3. **Comparison of Algorithm and Instructor Parameter Measurement (LUKE HI-ILS 3R APPROACH)**

Mission Element Number	Parameters Currently Programmed	Parameters Selected by IPs	Additions/Corrections
1	Airspeed, Radial, DME	Airspeed, Radial, DME	
2	Airspeed, Altitude	Airspeed, Altitude	
3	Airspeed, Altitude, DME	Airspeed, Altitude, DME	
4	Airspeed, Altitude, DME	Airspeed, Altitude, DME	
5	Airspeed, Altitude, DME	Airspeed, Altitude, DME	
6	Airspeed, Altitude, DME	Airspeed, Altitude, DME	
7	Airspeed, Altitude, DME	Airspeed, Altitude, DME	
8	Airspeed, Altitude	Airspeed, Altitude	
9	Airspeed, Altitude, Radial	Airspeed, Altitude, Radial	
10	Airspeed, Altitude, LOD[a]	Airspeed, Altitude, LOD[a]	
11	Airspeed, Altitude, Configuration	Airspeed, Altitude, Configuration	
12	Airspeed, LOD[a], GSD[b]	Airspeed, Altitude, Heading	Altitude
13	LOD[a]	Airspeed, Configuration, Heading	Airspeed, Configuration

[a]LOD includes both Radial and Heading Information.

[b]Glide Slope Deviation includes Altitude Information.

Table 4. **Comparison of Algorithm and Instructor Parameter Measurement (FLATIRON 03 DEPARTURE)**

Mission Element Number	Parameters Currently Programmed	Parameters Selected by IPs	Additions/Corrections
1	Heading, Channel	Heading, Airspeed, Configuration	Airspeed, Configuration
2	DME, Configuration	Airspeed, Altitude, DME	Airspeed, Altitude
3	Airspeed, Altitude, Heading	Airspeed, Altitude, Heading	
4	Airspeed, Altitude, Heading	Airspeed, Altitude, Heading	
5	Airspeed, Altitude, Radial[a]	Airspeed, Altitude, Heading	
6	Airspeed, Altitude, Radial[a]	Airspeed, Altitude, Radial[a]	
7	Airspeed, Altitude, Heading	Altitude, Heading, DME	DME
8	Airspeed, Channel, Heading	Altitude, Heading, Channel	Altitude
9	Airspeed, Altitude	Altitude, Radial[a], Heading	

[a]Radial includes Heading Information.

III. RESULTS

The results of the evaluation are presented and discussed according to two topical areas: (a) the importance of parameters scored and (b) the correlation of IP scores to algorithm scores, without modification.

Parameter Importance

Differences between parameters selected by the IPs and those currently in use by the scoring algorithms were negligible for the LUKE HI-ILS 3R and JAY HI-TACAN approaches and for the FLATIRON 03 departure. However, for the ILS RWY 11 approach, the IPs suggested changes in 28 percent of the possible parameters to be scored.

Correlation of IP and Computer Evaluations

A significant positive correlation existed for only one of the mission profiles–FLATIRON 03 departure. The approach profiles yielded nonexistent to negative correlations between IP and computer scores (see Table 5 and Appendix E for data).

Table 5. **Correlations Between Computer Scores and IP Scores**

Statistic	HI-ILS 3R APP	JAY-HI TAC APP	ILS 11 APP	FLATIRON 03 DEP
Pearson r =	–.806	–.102	–.533	.752
Spearman r =	–.908	.01	–.521	.639
N =	11	14	13	26

Discussion

An apparent discrepancy exists in the results. While the computer and the IP used similar parameters in evaluating a student's performance, a significant positive correlation existed for only one of the mission profiles – the departure profile. Discrepancies in the approach profiles could be corrected by improving simulator initializations for the approaches, and further research should be conducted before tolerance modifications are made.

Another reason which might account for the discrepancies in the approaches is the time period of data collection. Departures were obtained in the early periods of the study. However, due to organization of the syllabus and the resulting "press" on the student, full data on approaches could not be obtained until later stages of research.

The major problem of the study, however, was intrinsic to the computer. Departures were all initialized from the same location – the runway; therefore, the IP and the computer started their evaluations at identical points in space and time. However, not all students entered the approach patterns at identical locations. The IPs would often begin grading the approach before or after the computer's algorithms recognized the maneuver as such. Also, while the IP might find fault with the student's early maneuvers in the approach pattern, the computer would only grade the latter portion or vice versa. Thus, the IP evaluation could be lower/higher than that of the computer since grading may not have been initialized simultaneously at the same point in space. The FS, however, is capable of initializing the student to a location. If the student was initialized to a common point of entry in the approach patterns, the IP and computer evaluations would begin at identical points, and the discrepancy might be alleviated.

IV. RECOMMENDATIONS

To facilitate future studies on the parameters, the IP should receive a pre-experiment briefing, clearly outlining instructions and objectives of the study. Close coordination with the squadron should be required; and a research agreement drawn up that would allow maximum use of the FS by the squadrons and the research personnel. (This is now a standard procedure at AFHRL/FTO.)

Three areas of possible inquiry which might be useful in later research are (a) preprogrammed approaches and departures could be devised to determine IP relative grading levels for standardization, (b) weighted correction factors could be assigned that might subdue intra-IP discrepancy and yield a more meaningful correlation, and (c) prior to evaluating the computer program, a questionnaire, on which the parameter are to be scored, should be prepared and sent to the squadron areas. This last step would eliminate section 2 on the grade sheet, which requires a second IP evaluation two weeks after the first, and thus would remove additional work for the IP during the data collection phase of research.

In addition, future studies might consider modification and/or revision of the syllabus in order to receive greater experimental opportunities. It was impossible to obtain data from several periods due to heavy concentration on air work, etc. These lessons might be modified for departure and approach data collection purposes.

Implications

As currently programmed, the FS is qualified for use in the departure mission profile. This pilot study has left unanswered the question of whether discrepancies between IP and computer scores were due to parameter selection or to the difficulties encountered in scoring techniques and coordination. In either case, due to the unavailability of Standardization/Evaluation (STAN/EVAL) check pilots during this study, the FS should not be used for instrument checkrides based on computer scoring alone, at least not until extremely high positive correlations are obtained. The FS scoring program should aid students working together on "Buddy Rides" or it could be used in proficiency checks. The FS is a valuable tool that, with improvement, will aid both instructors and students.

REFERENCES

Cream, B.W., & Lambertson, D.C. *Functional integrated systems trainer: Technical design and operation.* AFHRL-TR-75-6(II), AD-A015 835. Wright-Patterson AFB, OH: Advanced Systems Division, Air Force Human Resources Laboratory, June 1975.

Ericksen, S.C. *A review of the literature on methods of measuring pilot proficiency.* Research Bulletin 52-25. Lackland AFB, TX: 1951.

Johnson, H.M., & Boots, M.L. *Analysis of ratings in the preliminary phase of the CAA training program.* Research Report No. 21, CAA Division of Research, 1943.

Knoop, P.A. Advanced instructional provisions and automated performance measurements. *Human Factors,* December 1973, **15**, 583–597.

Koonce, J.M. Effects of ground-based aircraft simulator motion conditions upon prediction of pilot proficiency. *Dissertation Abstracts International,* January 1975, **35**, 3557.

Long, G.E., & Varney, N.C. *The automated pilot aptitude measurement system.* Paper presented at the meeting of the Fifth Annual Symposium on Psychology in the Air Force, United States Air Force Academy, 1976.

Mahler, W.R., & Bennett, G.K. *Psychological studies of advanced naval air training; Analysis of flight performance ratings,* The Psychological Corporation. Technical Report SDC 99–1-2. Special Devices Center, 1950.

Obermayer, R.W., Vreuls, D., Muckler, F.A., Conway, E.J., & Fitzgerald, J.A. *Combat-ready crew performance measurement system: Final report.* AFHRL-TR-74-108(I), AD-B005 517L. Williams AFB, AZ: Flying Training Division, Air Force Human Resources Laboratory, December 1974.

TACR 60–2. *Aircrew standardization evaluation program.* 31 May 1977.

Waag, W.L., Eddowes, E.E., Fuller, J.H., Jr., & Fuller, R.R. *ASUPT automated objective performance measurement system.* AFHRL-TR-75-3, AD-A014 799. Williams AFB, AZ: Flying Training Division, Air Force Human Resources Laboratory, March 1975.

BIBLIOGRAPHY

Brown, J.E., Waag, W.L., & Eddowes, E.E. *USAF evaluation of an automated adaptive flight training system.* AFHRL-TR-75-55, AD-A018 612. Williams AFB, AZ: Flying Training Division, Air Force Human Resources Laboratory, October 1975.

Fitts, P.M., & Posner, M.I. *Human performance.* Belmont, CA: Brooks/Cole Publishing Company, 1967.

Henneman, R.H. Proficiency measures for fighter pilots at the operational level of training in the army air forces. *American Psychology,* 1946.

Koonce, J.M *Use of two axis tracking task in predicting pilot training success.* Paper presented at the meeting of the Fifth Annual Symposium on Psychology in Air Force, United States Air Force Academy, 1976.

Miller, N.E. (Ed.). *Psychological research on pilot training.* Research Report No. 8. AAF Aviation Psychology Program, 1947.

Prather, D.C. *Imagination as a flight simulator.* Paper presented at the meeting of the Third Annual Symposium on Psychology in the Air Force, United States Air Force Academy, 1974.

Schroder, J.H. A study of the proposed flying adaptability test (A.C.) in the course of physical examination of civilian aviators. *Journal of Aviation Medicine*, 1932.

Sinaiko, H.W., & Buckley, E.P. *Human factors in the design of systems.* Selected papers on human factors in the design and use of control systems. NRL Report 4996. Naval Research Laboratory, Washington, D.C., 1957.

Youtz, R.P., & Ericksen, S.C. *Analysis of the pilot's task.* Research Report No. 8. AAF Aviation Psychology Program, 1947.

REFERENCE NOTES

1. **Fessler, S.W.** *Pilot performance measurement: An analysis of the literature.* Unpublished manuscript, USAF Academy. 1977.

APPENDIX A: SAMPLE GRADING FORMS FOR APPROACHES AND DEPARTURE

The purpose of this study is to evaluate the scoring algorithms of the computer. We are trying to establish the amount of correlation between the computer's score of the student's performance and your evaluations. If the correlation is found to be low, the algorithms will be modified, based on inputs received from you. Therefore, we would appreciate your use of the following approach/departure patterns in your mission planning.

1. FLATIRON 03 Depature
2. JAY HI-TACAN APP
3. LUKE HI-ILS APP 03R
4. LUKE AUX #1-ILS 11

Grading System

1. Research grades will be completely anonymous; it is not necessary to enter either your name or the student's name on this form. Enter only mission number, date and time, rank, and experience.

2. Students should *not* see these grades. This frees you to assign grades freely, without having to worry about instilling either false confidence or discouragement among students.

3. In assigning grades of the maneuvers given, please grade as you would on any normal training mission for those maneuvers.

4. Grading will be done on a numerical scale. Definitions of the scores are provided on the Mission Grade Form. Scores will be marked on the line graph continuum according to your evaluation of the student's performance on the particular legs. Indicate your grade by placing a mark anywhere along the numerical grade scale. The scale has been expanded to allow for more discrimination in the middle range.

In order for us to make modifications to the computer's scoring program, we need your input for criteria to be graded. After the mission is completed, please check what you feel to be the *two* most important items to be considered for each leg. If you feel a criterion is important that is not listed, please write it in the space provided.

MISSION ELEMENTS

FLATIRON 03 DEP

Mission Elements	0 1 1.5 2 2.5 3 4	Speed	Altitude	Heading	Radial	DME	Channel	CFG	Other
1. Takeoff-030									
2. Left Turn-215									
3. Minimum altitude-4000									
4. Minimum altitude-6000									
5. Maintain radial-265									
6. Minimum Altitude-9000	0 1 1.5 2 2.5 3 4								
7. Right Turn-280									
8. Intercept radial-315									
9. Maintain radial-315									

GRADING CRITERIA

GRADE	DEFINITION
0	Performance indicates a lack of ability or knowledge.
1	Performance is safe but indicates limited proficiency. Makes errors of commission or omission.
2	Performance is essentially correct. Recognizes and corrects errors.
3	Performance is correct, efficient, skillful, and without hesitation.
4	Performance reflects an usually high degree of ability.

INDIVIDUAL TRAINING MISSION GRADE *(Grading Criteria On Reverse Side)*		Mission Number	Date	Time	
Student Rank	Total Simulator Time	Total A/C Time	F-15 A/C Time		

MISSION ELEMENTS HI-ILS 3R APP	0 1 1.5 2 2.5 3 4	Speed	Altitude	Heading	Radial	DME	Channel	Config	Other
1. Descent to ARC									
2. Rurn to ARC									
3. Maintain ARC to RAD 280									
4. Descent to RAD 255									
5. Cross RAD 255									
6. Descent to RAD 230									
7. Cross RAD 230									
8. Maintain ARC to LCZR									
9. Intercept LCZR	0 1 1.5 2 2.5 3 4								
10. LCZR to GPIP									
11. Glide Path Intercept Point									
12. Final Approach									
13. Missed APP or Landing									

AFHRL FORM JUN 77 135 (ONE-TIME) (Expires Sep 77)

INDIVIDUAL TRAINING MISSION GRADE
(Grading Criteria On Reverse Side)

Mission Number	Date	Time	

Student Rank	Total Simulator Time	Total A.C Time	F-15 A/C Time	

MISSION ELEMENTS

JAY HI-TAC APP

Mission Elements	0 1 1.5 2 2.5 3 4	Speed	Altitude	Heading	Radial	DME	Channel	Config	Other
1. Approaching LENNI									
2. Holding Pattern to DOODL									
3. Descent to ARC									
4. Maximum Altitude-16000									
5. Maintain ARC-16 DME									
6. Turn to Radial-312									
7. Descent to Final APP fix									
8. Final Approach fix-6000									
9. Descent to 5000	0 1 1.5 2 2.5 3 4								
10. Maintain Altitude-5000									
11. Final Approach									
12. Minimums-Missed APP Point									

AFHRL FORM JUN 77 135 (ONE-TIME) (Expires Sep 77)

MISSION ELEMENTS

ILS 11 APP

Mission elements	Scale (0 1 1.5 2 2.5 3 4)	Speed	Altitude	Heading	Radial	DME	Channel	CFG	Other
1. Maneuver to centerline	0 1 1.5 2 2.5 3 4								
2. Loc to glide path intercept pt.									
3. Glide path intercept point									
4. Final approach	0 1 1.5 2 2.5 3 4								
5. Minimums-1850									
6. Missed approach or landing									

GRADING CRITERIA

GRADE	DEFINITION
0	Performance indicates a lack of ability or knowledge.
1	Performance is safe but indicates limited proficiency. Makes errors of commission or omission.
2	Performance is essentially correct. Recognizes and corrects errors.
3	Performance is correct, efficient, skillful, and without hesitation.
4	Performance reflects an usually high degree of ability.

INDIVIDUAL TRAINING MISSION GRADE

(Grading Criteria On Reverse Side)

MISSION NUMBER	POSITION NUMBER	MISSION DURATION	DATE

NAME	AIRCRAFT NUMBER	TMS AIRCRAFT	INSTRUCTOR
		F-15 FLT SIM	

Simulator Conversion MISSION ELEMENTS	UNKNOWN	DANGEROUS	GRADE 0	GRADE 1	GRADE 2	GRADE 3	GRADE 4
1. MISSION PREPARATION							
2. PREFLIGHT	/	/	/	/	/	/	/
a. COCKPIT INTERIOR CHECK							
b. ENGINE START							
c. BEFORE TAXI CHECKS							
d. BEFORE T.O. CHECKS							
3. T.O. (MIL/AB)							
4. DEPARTURE	/	/	/	/	/	/	/
a. AIRSPEED CONTROL							
b. ALTITUDE CONTROL							
c. SID COMPLIANCE							
5. AIR WORK	/	/	/	/	/	/	/
a. BASIC AIRCRAFT CONTROL							
b. ADVANCED HANDLING							
c. UNUSUAL ATTITUDE RECOVERY							
d. STEEP TURNS							
6. RECOVERY	/	/	/	/	/	/	/
a. TACAN POINT-TO-POINT							
b. HOLDING							
c. TACAN PENETRATION/APPROACH							
d. GCA							
e. ILS							
f. MISSED APPROACH							
7. SYSTEMS OPERATION							
8. EMERGENCY PROCEDURES							
9. RADIO PROCEDURES							
10. FLIGHT MANAGEMENT							
14. AIRMANSHIP							
OVERALL GRADE							

REMARKS (Continue on reverse side)

Mission briefed and flown in accordance with Syllabus.

Supv

Inst

STUDENTS INITIALS	SIGNATURE OF INSTRUCTOR

AF FORM 1363 (PREVIOUS EDITION WILL BE USED)

MISSION ELEMENTS	U	D	0	1	2	3	4	REMARKS

GRADING CRITERIA

GRADE	DEFINITION
UNKNOWN	Performance not observed or the element was not performed.
DANGEROUS	Performance was unsafe.
0	Performance indicates a lack of ability or knowledge.
1	Performance is safe but indicates limited proficiency. Makes errors of commission or omission.
2	Performance is essentially correct. Recognizes and corrects errors.
3	Performance is correct, efficient, skillful, and without hesitation.
4	Performance reflects an usually high degree of ability.

APPENDIX B: SAMPLE APPROACH PLATE WITH NUMBERED SEGMENTS

This form is an opportunity for us to receive detailed input on your feelings of the criticality of certain criteria during legs of approaches and departures. During our research this summer we found a positive correlation to exist between your grades of a student's performance and the computer's evaluations in terms of ranking the students. However, we feel the algorithms could be significantly improved by modifying existing tolerance limits or by having the computer evaluate different criteria.

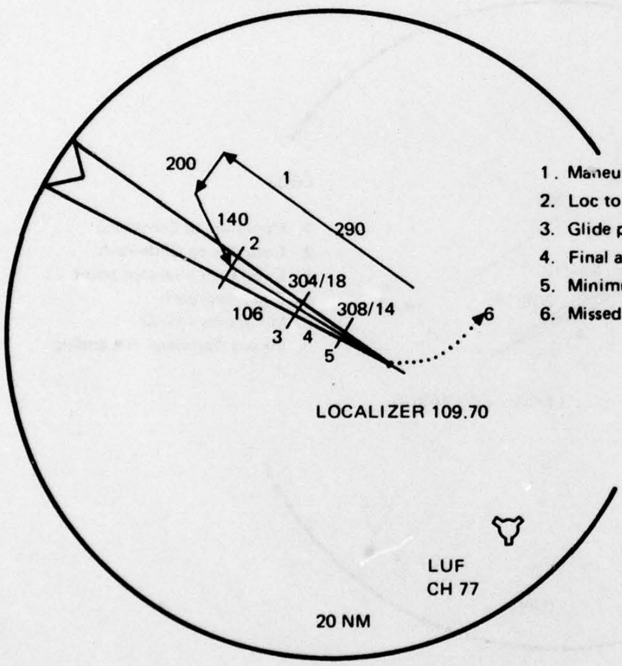
Numbered diagrams of the approaches and departures are presented, the numeral marking the end of a leg. The number is briefly described and blocks labeled with currently used criteria are listed on the side. In these blocks please check what you feel to be the three most critical criteria for each leg. However, if you feel another criterion is more important, place the number of it from the list below in the "other" block. This list represents all the evaluations the computer is capable of.

1. SPEED
2. ALTITUDE
3. HEADING
4. RATE OF CLIMB
5. RATE OF TURN
6. ANGLE OF ATTACK
7. ROLL RATE
8. GLIDE SLOPE DEVIATION
9. LOCALIZER DEVIATION
10. TACAN RADIAL
11. TACAN RANGE
12. TACAN CHANNEL
13. CONFIGURATION
14. LATITUDE
15. LONGITUDE
16. GROSS WEIGHT
17. TOUCHDOWN

When you have completed this questionnaire, please leave it at your Squadron Operations Desk.

We from the Air Force Academy sincerely appreciate your cooperation throughout the summer. In the next three weeks we will complete our data collection as well as receive your inputs on this form. We realize that our collection presents an inconvencience at times, but hope that we can continue to receive your support in an effort to make your job as an Instructor Pilot a little easier.

LUKE AUX1
ILS 11



1. Maneuver to centerline
2. Loc to glide path intercept pt.
3. Glide path intercept point
4. Final approach
5. Minimums-1850
6. Missed approach or landing

Speed	Altitude	Heading	Radial	DME	Channel	Config	Other

APPENDIX C: SAMPLE APPROACH PLATE WITH PARAMETERS LISTED

APP CON
256.90
TOWER
289.60
GND CON

LUKE AUX1
ILS 11

P TIME 0 31 01

1 M-SPD 245.25 +-17.25
M-ALT 3500 +1700
M-CHL 77
T-LOD 0 +-1.219

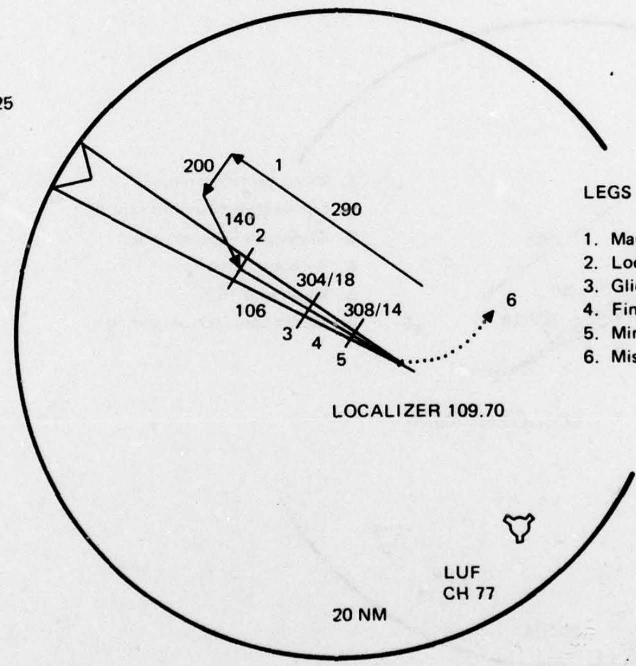
2 M-ALT 2800 +1100
M-LOD 0 +-1.219
T-GSD 0 +-0.5
ILS LOCKON

3 M-SPD 152.5 +-7.5
M-LOD 0 +-1.219
M-CFG 6 +1
T-GSD 0 +-0.5

4 M-SPD 152.5 +-7.5
M-GSD 0 +-0.219
M-LOD 0 +-1.219
T-DME 13.1 +-1

5 M-ALT 1875 +-75
M-GSD 0 +-0.219
M-LOD 0 +-1.219
T-DME 13.1 +-1

6 M-LOD 0 +-1.219
LAST LEG



LEGS

1. Maneuver to Centerline
2. Localizer to Glide Path
3. Glide Path intercept point
4. Final Approach
5. Minimums - 1850'
6. Missed Approach or Landing

APPENDIX D: SAMPLE COMPUTER FLIGHT SCORING READOUT

FLIGHT SCORING DEPARTURE

AIRFIELD: LUKE AFB FLATIRON GLAD

Leg No.	P Time	DME	Radial	Parameter	Max. Val	Time Out
1	0:20:57	0.7	86	HDG	24.9	2
1	0:21:01	0.8	71	HDG	24.5	4
1	0:21:27	2.8	37	HDG	24.2	2
3	0:22:12	7.8	323	SPD	386.4	23
4	0:23:36	7.9	299	SPD	390.0	34
4	0:24:01	9.3	278	HDG	233.2	9
5	0:24:11	10.2	272	SPD	389.3	36
6	0:24:48	15.2	267	SPD	379.6	21
7	0:26:33	29.3	265	SPD	368.2	11
7	0:26:46	31.1	265	SPD	373.6	4
7	0:26:52	23.8	265	HDG	254.3	58

SCORE 69

APPENDIX E: DATA FOR CORRELATION ANALYSIS

IP Score	Rank	Computer Score	Rank
FLATIRON DEPARTURE N = 26			
3.00	1	96	2.5
2.94	2.5	96	2.5
2.94	2.5	98	1
2.89	4.5	89	7.5
2.89	4.5	81	11.5
2.85	6	89	7.5
2.78	7	87	9
2.77	8	59	21
2.72	9	73	17
2.66	10	76	15
2.61	11.5	92	5
2.61	11.5	81	11.5
2.55	13	59	21
2.50	15	75	16
2.50	15	70	18
2.50	15	42	24
2.44	17	95	4
2.33	18	83	10
2.16	19	79	13
2.14	20	91	6
1.97	21	77	14
1.83	22	59	21
1.77	23	60	19
1.71	24.5	56	23
1.71	24.5	10	25
1.00	26	0	26
JAY HI TACAN N = 14			
2.83	1	100	1
2.75	2.5	57	6.5
2.75	2.5	26	11
2.52	4	18	12
2.38	5	37	9
2.25	6	69	2
2.22	7	0	14
2.13	8	66	3
1.95	9	57	6.5
1.93	10	16	13
1.92	11	52	8
1.83	12	33	10
1.77	13.5	61	4
1.77	13.5	60	5

Appendix E (*Continued*)

IP Score	Rank	Computer Score	Rank
HI ILS N = 11			
2.77	1	59	11
2.72	2	73	9
2.66	3	76	7
2.61	4	81	5
2.50	5.5	75	8
2.50	5.5	71	10
2.44	7	95	3
2.16	8	79	6
2.14	9	91	4
2.00	10	100	1
1.77	11	99	2
AUX ILS N = 13			
2.91	1	28	12
2.50	2.5	12	13
2.50	2.5	88	6
2.38	4	85	7.5
2.22	5	91	4.5
2.00	7	91	4.5
2.00	7	85	7.5
2.00	7	61	10
1.96	9	92	3
1.94	10	38	11
1.67	11.5	76	9
1.67	11.5	96	1
1.50	13	93	2

☆U.S. GOVERNMENT PRINTING OFFICE:1979- 671-056/28